An engineered E. coli strain for direct in vivo fluorination

A EUROPEAN JOURNAL OF CHEMICAL BIOLOGY

CHEMBIOCHEM

SYNTHETIC BIOLOGY & BIO-NANOTECHNOLOGY

Accepted Article

Title: An engineered E. coli strain for direct in vivo fluorination

Authors: Konstantinos I. Markakis, Phillip T. Lowe, Liam Davison-Gates, David O'Hagan, Susan J. Rosser, and Alistair Elfick

To be cited as: *ChemBioChem* 10.1002/cbic.202000051

Link to VoR: http://dx.doi.org/10.1002/cbic.202000051

A Journal of

ChemPubSoc Europe

WILEY-VCH

www.chembiochem.org

WILEY-VCH

COMMUNICATION

An engineered *E. coli* strain for direct *in vivo* fluorination

Konstantinos Markakis*[a], Phillip T. Lowe[b], Liam Davison-Gates[a], David O'Hagan[b], Susan J.Rosser[c] and Alistair Elfick*[a]

[a] Dr. K. Markakis*, L. Davidson-Gates, Prof. A. Elfick*
Institute for Bioengineering
School of Engineering, University of Edinburgh
Faraday Building, King's Buildings
Colin Maclaurin Road
Edinburgh EH9 3DW
United Kingdom
E-mail: Alistair.Elfick@ed.ac.uk, K.Markakis@ed.ac.uk

[b] Dr. P. T. Lowe, Prof. D. O'Hagan
School of Chemistry
University of St. Andrews
Purdie Building
North Haugh
St. Andrews
United Kingdom

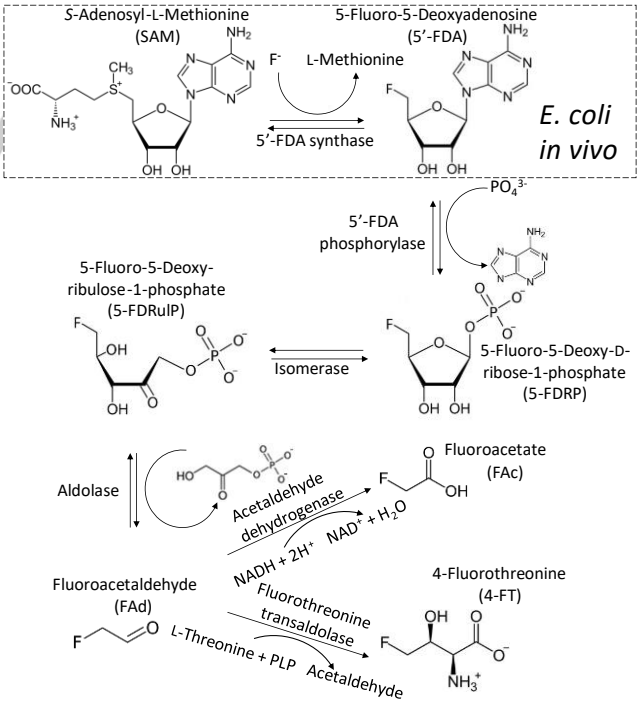
[c] Prof. S. Rosser
School of Biological Sciences, The University of Edinburgh
Roger Land Building
Alexander Crum Brown Road
The King's Buildings
Edinburgh EH9 3FF
United Kingdom

Supporting information for this article is given via a link at the end of the document.

Abstract: *Selectively fluorinated compounds are found frequently in pharmaceutical and agrochemical products where currently 25-30% of optimized compounds emerge from development containing at least one fluorine atom. There are many methods for site specific introduction of fluorine, but all are chemical and they often use environmentally challenging reagents. Biochemical processes for C-F bond formation are attractive but they are extremely rare. In this work the fluorinase enzyme, originally identified from the actinomycete bacterium* S. cattleya*, is engineered into* E. coli *in a manner where the organism is able to produce 5'-fluorodeoxyadenosine (5'-FDA) from S-adenosyl-L-methionine (SAM) and fluoride in live* E. coli *cells. Success required the introduction of a SAM transporter and deletion of the endogenous fluoride efflux capacity, in order to generate an* E. coli *host which has potential for future engineering of more elaborate fluorometabolites.*

Fluorine is the most abundant halogen found in the Earth's crust, but despite its geological dominance among the halogens, the concentration of soluble fluoride in the oceans and surface water is very low and its bioavailability is restricted.[1] Additionally, a fluoride ion is not readily oxidised like the other halogens e.g. by haloperoxidases. Moreover, it is the most strongly hydrated halide, thus catalytic processes need to overcome considerable kinetic barriers to achieve reactivity. Therefore, the occurrence of fluorometabolites is extremely rare and a biotechnology associated with organofluorine production from fluoride is essentially absent.[2] This can be contrasted with the very large anthropogenic focus on fluorine for the fine tuning of properties associated with the industrial development of performance chemicals. As a result, a rich and wide-ranging chemistry has emerged, associated with the chemical incorporation of fluorine in performance organic molecules ranging from materials to bioactives.[3]

The first fluorometabolite to be identified in nature (1944)[4] was fluoroacetate (FAc), an acute toxin which has subsequently been shown to be produced by a range of tropical and subtropical plants.[5, 6] Its origin in plants remains unknown, however the biosynthesis of FAc and co-produced 4-fluorothreonine (4-FT) in bacteria has been elucidated in *Streptomyces cattleya*.[7, 8] This actinomycete contains six genes encoding enzymes which process inorganic fluoride and SAM to FAc and 4-FT (Scheme 1). The pathway itself has been reconstituted *in vitro* by combining the over-produced enzymes into an NMR tube.[9] The first



Scheme 1. The 4-FT - FAc pathway from *S. cattleya.* The reaction catalyzed by 5'-FDA synthase (fluorinase) generates 5'-FDA from SAM.[10, 11] This is the first step in fluorometabolite biosynthesis, a rare process which has been identified only in a small number of actinomycete bacteria.[12]

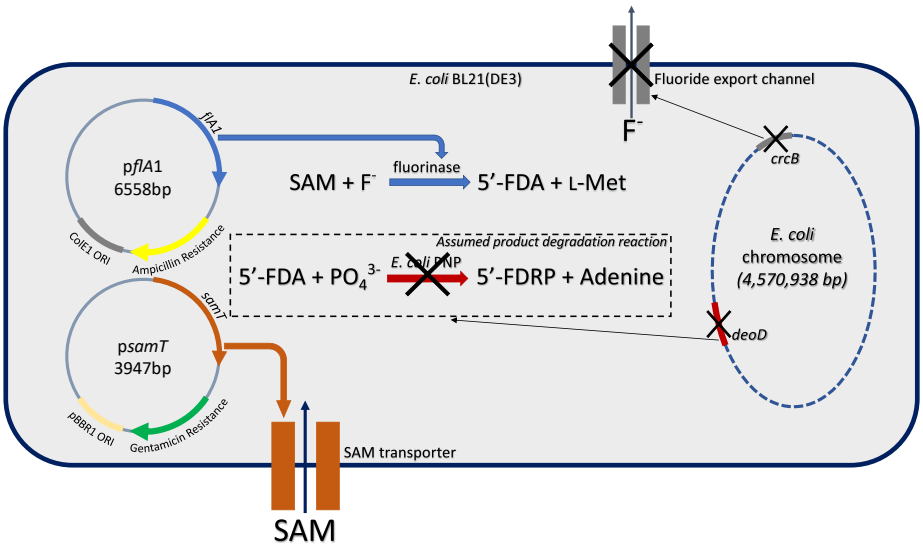


Figure 1. Schematic of the variables in the engineered *E.coli* BL21(DE3) cells showing plasmids carrying the fluorinase (*flA*1) and SAM transporter (*samT*) genes, as well as the chromosomal knockouts of the fluoride ion channel (*crcB*) and purine nucleoside phosphorylase (*deoD*).

committed enzyme in the 4-FT and FAc pathway is the fluorinase, which catalyzes C-F bond formation during the conversion of SAM to 5'-fluoro-5'-deoxyadenosine (5'-FDA)[10, 11] and this became the target enzyme for *in vivo* expression in *E. coli* as illustrated in Scheme 1.

Subsequent to the original fluorinase isolation from *S. cattleya*, four more variants have been functionally demonstrated after their identification by genome mining. All genes exhibited high sequence homology (80%-87%) with the original gene from *S. cattleya* and all corresponding enzymes had a similar catalytic efficiency.[12] These included a fluorinase from the marine bacterium *Streptomyces xinghaiensis*, which was able to produce FAc in a sea salt medium.[13, 14] As more bacterial genome sequences are deposited in public databases, it seems likely that more fluorinases will emerge. A current BLAST search identified two more open reading frames (ORFs) which are predicted to encode for fluorinases with a sequence identity of ~80%, (from *Actinopolyspora mzabensis* and *Amycolatopsis sp.* CA-128772 genomes).

In this work, a rational strategy for developing an engineering platform for *in vivo* fluorination based on *E. coli* BL21(DE3) was explored, with a specific focus to establish 5'-FDA production. This particular strain was used as it has already been utilized in fluorinase over-expression. A particular concern was the well-described ability of fluoride to inhibit growth in bacteria.[15] Fluoride enters the cell by passive diffusion across the membrane as neutral hydrogen fluoride (HF), where it then dissociates to fluoride ions.[16] It has been reported that the presence of cellular fluoride in the millimolar range inhibits some key enzymes involved in glucose metabolism, such as enolase[17] and pyrophosphatase.[18] Several studies have revealed a wide variety of proteins that play a role in managing high fluoride ion concentrations in bacterial or eukaryotic hosts.[15, 19-21] These are transmembrane channels that serve as pumps expelling fluoride which has entered passively as HF. *E. coli* employs such a channel protein termed CrcB which it uses to achieve fluoride resistance.[15] Mutant *E. coli* strains with the *crcB* gene deleted, exhibit a 200-fold greater sensitivity to extracellular fluoride relative to wild type. The mechanism of action of CrcB was evidenced by the fact that the intracellular concentration of fluoride in a Δ*crcB* strain remained at much higher levels compared to the wild type, under the same extracellular fluoride concentration.[15] Additionally, passive HF trans-membrane ingress is a pH-dependent process.[16] A study in *S. cattleya* has indicated that fluoride uptake - and therefore intracellular availability - directly influences *in vivo* fluorometabolite production by triggering the upregulation of key biosynthetic enzymes.[22] This has been attributed to a fluoride uptake mechanism in *S. cattleya*, although such a mechanism remains to be confirmed. Alternatively, the absence of a fluoride efflux channel in *S. cattleya* would be sufficient to allow fluoride to accumulate passively, consistent with the pH-dependence of fluoride uptake and the growth retardation of *S. cattleya* above 2 mM added fluoride [23]. In other fluoride channel-expressing bacteria, such as *E. coli*, millimolar fluoride supplied in the growth medium has little effect on growth. In this context, deletion of the *crcB* gene was explored to raise intracellular fluoride concentrations, and counter the relatively poor affinity of fluoride ion for the fluorinase (millimolar level K_m for fluoride). [11, 12]

A second modification addressed deletion of the *deoD* gene encoding a purine nucleoside phosphorylase (PNP) in *E. coli*. This enzyme can cleave purine bases including adenosine from purine nucleosides to generate ribose-1-phosphate.[24, 25] A PNP in *S. cattleya* catalyzes the second step of the fluorometabolite pathway and converts 5'-FDA to yield 5-fluoro-5-deoxy-D-ribose 1-phosphate (5-FDRP).[26] The logic behind *deoD* deletion was to remove this PNP activity and suppress any adventitious 5'-FDA degradation. Such a consideration was noted in a recent study exploring directed evolution of the fluorinase involving *E. coli* cell-free extract incubations.[27]

A third modification involved the expression of a gene encoding a transmembrane SAM transporter in order to increase the pool of intracellular SAM for the fluorinase reaction. SAM transporters are found in symbiotic organelles such as mitochondria[28, 29] and plastids.[30] Some intracellular parasites, including *Rickettsia prowazekii* RP076 and *Chlamydia trachomatis* contain SAM transporter genes. These have been expressed in otherwise non-viable *E. coli* strains, with the gene encoding for the SAM producing enzyme (*metK*) deleted. Expression of the transporter genes with external supplementation of SAM rescued growth in these strains.[31, 32] We chose the SAM transporter from *R. prowazekii* (hereafter denoted SAMT) because it is an energy-dependent channel protein transporting SAM only from the extracellular to the intracellular environment contrary to the *Chlamydial* transporter which exchanges SAM for *S*-adenosyl-L-homocysteine (SAH).

10.1002/cbic.202000051

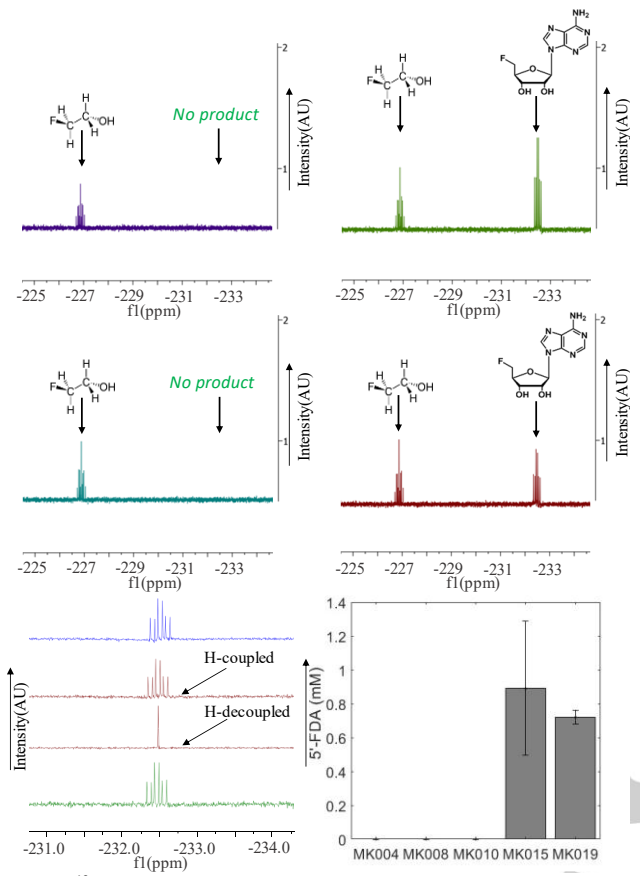


Figure 2. ^{19}F-NMR spectra of lysates from strains MK004 – BL21(DE3) + p*flA*1 **(Top-left)**, MK010 – Δ*crcB* + p*flA*1 **(Middle-left)**, MK015 – Δ*crcB* + p*flA*1 + p*samT* **(Top-right)**, MK019 – Δ*crcB*Δ*deoD* + p*flA*1 + p*samT* **(Middle-right)**. 2-Fluoroethanol was used as an internal reference (-226.8 ppm) and 5'-FDA is detected at -232.5 ppm. The multiplicities arise from couplings between ^{19}F and ^{1}H nuclei. **Bottom-left:** Zoomed ^{19}F-NMR spectrum of the fluorometabolite peak. 5'-FDA reference (blue), cell lysate of strain MK015 spiked with synthetic 5'-FDA (red, H-coupled and H-decoupled) and cell lysate of the same strain non-spiked (green). **Bottom-right:** Production yields in different strains as measured from 2 biological replicates. The computed molarity corresponds to the intracellular concentration of 5'-FDA.

Moreover, optimal growth of the Δ*metK* strain was observed with a much lower molarity of supplemented SAM in the growth medium.

Figure 1 illustrates the key modifications of the engineered *E. coli* and Table S2 summarizes the engineered *E. coli* strains and comparative incubations carried out in order to identify the impact that each modification had on fluorometabolite production. The fluorinase gene was adopted from *Streptomyces sp*. *MA37* (*flA*1) which was characterized as the most efficient fluorinase based on enzyme kinetics [12] and was inserted in the expression plasmid after codon optimization for *E. coli* (gene sequence in SI). Both deletions were achieved by suicide plasmid integrations in the chromosome and homologous recombination as originally described.[33] Incubations were supplemented with 2 mM KF and 500 µM SAM. For a more detailed description see *Supporting Information.*

It is noteworthy that fluorinase over-expression in all tested strains seemed to stimulate growth to a higher bacterial mass during stationary phase. In order to further explore this effect, strains BL21(DE3) (no plasmids), MK007 (BL21(DE3) with negative control plasmid pET(-)) and MK004 (BL21(DE3) with fluorinase expression plasmid, p*flA*1) were grown in 96-well plates and their growth was monitored. Results are presented in Figure S9. Indeed, higher growth is observed in the BL21(DE3) strains expressing the fluorinase when compared to BL21(DE3) strains free of plasmids, negative control (MK007 – pET(-)) or basal expression (MK004 without IPTG). Generally, protein over-expression will impede growth even in the case of over-production of an inert protein.[34] To explore if this effect was due to increased levels of intracellular protein accumulation due to over-expression, varying levels of induced expression of the fluorinase (IPTG 0.02 mM to 0.08 mM) were included. The indifference of stationary OD levels between different expression levels confirms that the higher stationary optical density (OD) observed is not due to protein accumulation. The origin of this effect is unclear but it bodes well for future development as recombinant expression of heterologous proteins can often lead to genetic instability of the host, due to toxicity or added metabolic burden.[35]

Initially, strains including either all modifications or all but PNP deletion (thought as most promising for achieving *in vivo* fluorination) were tested, specifically strains MK015 (Δ*crcB* + p*flA*1 + p*samT*) and MK019 (Δ*crcB*Δ*deoD* + p*flA*1 + p*samT*) with their negative control counterparts MK016 and MK020. Lysates and supernatants from two replicates were assayed by HPLC and also ^{19}F-NMR, which directly reports fluorinated products. The lysates in both of fluorinase containing strains MK015 and MK019 showed clear ^{19}F-NMR signals associated with 5'-FDA as the only organo-fluorine product whereas the corresponding negative controls (MK016 and MK020) without the *flA*1 gene, did not.

A series of experiments (see SI for details) explored incubations of engineered *E. coli's* with and without *samT* expression and *crcB* gene active or deleted. Cell extracts were evaluated by ^{19}F-NMR (Figure 2). No 5'-FDA was detected in MK004 and MK010, both of which express *flA*1 but not *samT*, with MK010 having an inactivated fluoride efflux channel. This is contrasted with 5'-FDA production in MK015 (PNP intact) and MK019 (PNP inactive) both of which express *samT* but have their efflux channels inactivated. The stability of 5'-FDA in these strains suggests that it is not degraded *in vivo* by the PNP or other potential metabolizing enzymes such as adenosine deaminase as no additional fluorometabolites were identified by ^{19}F-NMR. It should be noted that in cell-free extracts of *S. cattleya* 5'- FDA is metabolised to 5'-FDI by an endogenous deaminase, a process that does not occur in whole cell incubations, suggesting some level of compartmentalisation, which may be the case in *E. Coli* as well.[24] The necessity for both fluoride channel deletion and *samT* expression to achieve 5'-FDA production was recognised with strain MK008 (Figure S10). This strain which has an active fluoride efflux channel, as well as *flA*1 and *samT* genes, was unable to produce 5'-FDA. From this data it would appear that each modification on its own is insufficient to enable 5'-FDA production in the cells. Apart from expression of *flA1,* both deletion of *crcB* and expression of *samT* are required for fluorometabolite synthesis in *E. coli.*

The engineered *E. coli* developed here can convert fluoride to an organofluorine metabolite (5'-FDA) *in vivo*. The only other fluorometabolite produced by heterologous expression of the fluorinase gene from *S. cattleya* (*flA*) is fluorosalinosporamide which was induced in the marine organism *Salinospora tropica.* In that case, *flA* was inserted directly into a target gene to disrupt a highly homologous chlorinase gene within the (chloro)salinosporamide gene cluster. This resulted in the production of the fluoroanalogue.[36] Although a notable outcome, the gene swap was very conservative and replaced a like-for-like gene in a biosynthetic cluster. Also, the engineered *S. tropica* was extremely susceptible to fluoride ion toxicity which compromised the ability to carry out practical biotransformations. The ability of this engineered *E. coli* to produce 5'-FDA by manipulating intracellular fluoride ion and SAM concentrations, offers a first

important step towards the successful engineering of fluorometabolites in industrial microorganisms.

Acknowledgements

This work was funded by the Industrial Biotechnology Innovation Centre (IBioIC) with support from GlaxoSmithKline, and also the EU Horizon 2020 (Sinfonia consortia). We thank Dr Aitor De Las Heras and Dr Vlastimil Šršen for advice and Prof. David Wood, University of South Alabama, for an *E. coli* strain harbouring the *R. prowazekii* SAM transporter.

Keywords: fluorinase • halogenation • *E. coli* • SAM transporter • fluoride channel

[1] J. Emsley, Oxford University Press, **2001**,

[2] H. Deng, D. O'Hagan, *Chem. Rev*., **2015**, *115*, 634-649.

[3] (a) N. A. Meanwell, *J. Med. Chem.,* **2018**, *61*, 5822-5880; (b) H. Mei, J. Han S. Fustero, M. Medio‐Simon, D. M. Sedgwick, C. Santi, R. Ruzziconi, V. A. Soloshonok, *Chem. Eur. J.,* **2019**, *25*, 11797 – 11819.

[4] J. S. C. Marais, Onderstepoort *J. Vet. Sci. Anim. Ind*., **1944**, 20, 67-73.

[5] P. Oelrichs, T. Mc*Ewan, Nature, **1961**,190, 808-809.*

[6] D. O'Hagan, R. Perry, J.M. Lock, J.J.M. Meyer, L. Dasaradhi, J.T.G. Hamilton, D.B. Harper, *Phytochemistry*, **1993**, *33*, 1043-1046.

[7] M. Sanada, T. Miyano, S. Iwadare, J. M. Williamson, B. H. Arison, J. L. Smith, A. W. Douglas, J. M. Liesch, E. Inamine, *J. Antibiotics,* **1986**, *39*, 259-265.

[8] K. K. J. Chan, D. O'Hagan, *Methods Enzymol*., **2012**, *516*, 219-235.

[9] H. Deng, S. M. Cross, R. P. McGlinchey, J. T. Hamilton, D. O'Hagan, *Chem. Biol*., **2008**,*15*, 1268-1276.

[10] D. O'Hagan, C. Schaffrath, S. L. Cobb, J. T. Hamilton, C. D. Murphy, *Nature*, **2002**, *416*, 279.

[11] C. Dong, F. L. Huang, H. Deng, C. Schaffrath, J. B. Spencer, D. O'Hagan, J. H. Naismith, *Nature,* **2004**, *427*, 561 – 565.

[12] H. Deng, L. Ma, N. Bandaranayaka, Z. Qin, G. Mann, K. Kyeremeh, Y. Yu, T. Shepherd, J. H. Naismith, D. O'Hagan, *ChemBioChem.,* **2014**, *15*, 364-368.

[13] S. Huang, L. Ma, M. H. Tong, Y. Yu, D. O'Hagan, H. Deng, *Org. Biomol. Chem.,* **2014**, *12*, 4828-4831.

[14] L. Ma, Y. Li, L. Meng, H. Deng, Y. Li, Q. Zhang, A. Diao, *RSC Adv*., **2016**, *6*, 27047-27051.

[15] J. L. Baker, N. Sudarsan, Z. Weinberg, A. Roth, R. B. Stockbridge, R. R. Breaker, *Science*, **2012**, *335*,233-235.

[16] C. Ji, R. B. Stockbridge, C. Miller, *J. Gen. Physiol.,* **2014**, 144, 257-261.

[17] J. Qin, G. Chai, J. M. Brewer, L. L. Lovelace, L. Lebioda*, Biochemistry,* **2006**, *45*, 793-800.

[18] V. R. Samygina, V. M. Moiseev, E. V. Rodina, N. N. Vorobeyeva, A. N. Popov, S. A. Kurilova, T. I. Nazarova, S. M. Avaeva, H. D. Bartunik, *J. Mol. Biol.,* **2007**, **366**, 1305-1317.

[19] D. L. Turman, A. Z. Cheloff, A. D. Corrado, J. T. Nathanson, C. Miller, *Biochemistry,* **2018**, *57*, 1212 – 1218.

[20] S. Li, K. D. Smith, J. H. Davis, P. B. Gordon, R. R. Breaker, S. A. Strobel, *Proc. Nat. Acad. Sci.,* **2013**, *110*, 19018-19023.

[21] R. B. Stockbridge, H.-H. Lim, R. Otten, C. Williams, T. Shane, Z. Weinberg, C. Miller, *Proc. Nat. Acad. Sci.,* **2012**, 109,15289-15294.

[22] M. C. Walker, M. Wen, A. M. Weeks, M. C. Y. Chang, *ACS Chem. Biol*., **2012**, 7, 1576 – 1585.

[23] K.A. Reid, R.D. Bowden, L. Dasaradhi, M.R. Amin, D.B. Harper, *Microbiol*., **1995**, *141*, 1385-1393.

[24] G. W. Koszalka, J. Vanhooke, S. Short, W. W. Hall, *J. Bacteriol*., **1988**, *170*, 3493-3498.

[25] E. M. Bennett, C. Li, P. W. Allan, W. B. Parker, S. E. Ealick, *J. Biol. Chem.,* **2003**, *278*, 47110-47118.

[26] F. Huang, S. F. Haydock, D. Spiteller, T. Mironenko, T-L. Li, D. O'Hagan, P. F. Leadlay, J. B. Spencer, *Chem. Biol.,* **2006**, *13*, 475-484.

[27] H. Yeo, W.L. Sun, Y.H. Lim, X. Chew, D.J. Smith, B. Xue, K.P. Chan, R.C. Robinson, E.G. Robins, H. Zhao, E.L. Ang, *Angew. Chemie*, **2016**, *128*, 14489-14492.

[28] C. Marobbio, G. Agrimi, F. Lasorsa, F. Palmieri, *EMBO J*., **2003**, *22*, 5975-5982.

[29] G. Agrimi, M. Di Noia, C. Marobbio, G. Fiermonte, F. Lasorsa, F. Palmieri, *Biochemical J.,* **2004**, *379*, 183-190.

[30] F. Bouvier, N. Linka, J.-C. Isner, J. Mutterer, A. P. Weber, B. Camara, Plant Cell, **2006**,*18*, 3088-3105.

[31] A. M. Tucker, H. H. Winkler, L. O. Driskell, D. O. Wood, *J. Bacteriol.,* **2003**, *185*, 3031-3035.

[32] R. Binet, R. E. Fernandez, D. J. Fisher, A. T. Maurelli, *mBio* **2011**, *2*, e00051-11.

[33] G. Pósfai, V. Kolisnychenko, Z. Bereczki, F. R. Blattner, *Nuc. Acid Res.,* **1999**, *27*, 4409 - 4415.

[34] A. Das, S. Biswas, M. Biswas, *Open Microbial. J,* **2018**, *12*, p.107.

[35] K.E. Tyo, P.K. Ajikumar, G. Stephanopoulos, *Nature Biotechnol.*, **2009**, *27*, 760-765.

[36] A. S. Eustáquio, D. O'Hagan, B. S. Moore, *J. Nat. Prod*., **2010**, *73*, 378-382.

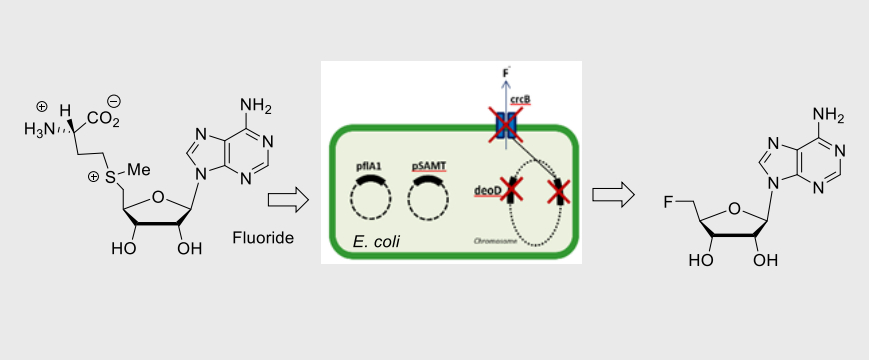
10.1002/cbic.202000051

WILEY-VCH

COMMUNICATION

Entry for the Table of Contents (Please choose one layout)

COMMUNICATION



Fluorinated organic molecules are of great value both for agriculture and medicine. However, fluorine chemistry is expensive and harmful to the environment. Natural products are at a premium and are only produced by a handful of organisms. For the first time, an engineered *E. coli* is presented with the capacity to produce a naturally occurring fluorinated molecule.

Konstantinos Markakis, Phillip T. Lowe, Liam Davison-Gates, David O'Hagan, Susan J.Rosser and Alistair Elfick**

Page No. – Page No.

An engineered *E. coli* strain for direct *in vivo* fluorination